

## **REMARKS**

Claims 1-12, 14-37 and 39-42 were presented for examination and were rejected. Applicants thank the Examiner for examination of the claims pending in this application and addresses Examiner's comments below.

Reconsideration of the application in view of the above amendments and the following remarks is respectfully requested.

In a final office action mailed October 11, 2006, the Examiner rejected the claims 1-3, 6-12, 14-18, 35-37 and 39 under 35 U.S.C. § 103(a) as being unpatentable over Rubin et al ("Rubin," US 2002/0099552) in view of Heck et al ("Heck" "A Survey of Web Annotation Systems"). The Examiner also rejected claims 4, 5, 19-34 and 40-42 under 35 U.S.C. § 103(a) as being unpatentable over Rubin et al ("Rubin," US 2002/0099552) and Heck et al ("Heck" "A Survey of Web Annotation Systems") in view of Mitchell et al ("Mitchell," US 5,857,099).

Applicants have amended claims 1, 9, 11, 12, 14, 26 and 28. Applicants have canceled claims 4, 7, 8, 10, 13, 15, 17-25 and 38.

Claim 1 is representative of the independent claims and has been amended to recite:

"a storage device for storing a plurality of different visual notations each comprising

a text or a graphic image and for storing a plurality of corresponding audio signals;

a direct annotation creation module coupled to receive the audio signal from the audio

input device and to receive a reference to a location within an image on the

display device, the direct annotation creation module, in response to receiving

the audio signal and the reference to the location within the image, automatically creating an annotation object, independent from the image, that associates the input audio signal, the location and one of the plurality of different visual notations; and

an audio vocabulary comparison module coupled to the audio input device, the audio vocabulary storage and the direct annotation creation module, the audio vocabulary comparison module receiving audio input and finding a corresponding one of the plurality of different visual notations that matches the audio input.”

The claimed invention is directed to the direct multi-modal annotation of objects. This is reflected in the language “direct annotation creation module.” The process of labeling an annotation is one of the key differences between the claimed invention and prior art. This is set forth in the claim language “a plurality of different visual notations each comprising a text or a graphic image.” In particular, the claimed invention compares input with stored exemplars, labels the location of input with stored exemplars and then labels/links the input with the corresponding exemplar label. This is set forth in the claim language “the audio vocabulary comparison module receiving audio input and finding a corresponding one of the plurality of different visual notations;” “automatically creating an annotation object, independent from the image, that associates the input audio signal, the location and one of the plurality of different visual notations,” respectively. Furthermore, the dependent claims provide for automatic creation of a new exemplar and associating a label if there is no sufficiently good match for the input.

The claimed invention is advantageous in a number of respects none of which are shown, described, taught or suggested by Rubin-Heck- Mitchell alone or in combination.

First, the claimed invention compares the input to a stored vocabulary with the audio vocabulary comparison module. This is not full-on speech recognition, but comparison between an input and stored exemplars. Since there are likely to be only a limited number of labels that are reused, this is much faster and accurate than traditional speech recognition. Furthermore, even if Rubin-Heck- Mitchell were combined, they would not yield the claimed invention where there is a relationship between the label and the image/position and the audio. There is no teaching or suggestion of a three way relationship in the Rubin-Heck- Mitchell combination.

Second, the claimed invention provides distinct labels or visual notations each comprising a text or a graphic image. Assuming a matching exemplar is found, the claimed invention labels the input with the label associated with that exemplar. For example, particular labels may be text or phases that describe a portion of the image such as "Uncle Fred" for one location and "Auntie June" for another location. This differs dramatically from the cited art where the label is simply the same generic audio icon for all annotations, or simply differences in color.

Third, the claimed invention allows for the easy creation of new labels since the vocabulary is may limited. If a good match for an annotation is not found, the user is given the option to either create a new exemplar (and label) or pick an existing exemplar. In the second case, the audio input is used as and additional training point for that exemplar (meaning the accuracy in recognizing the input will increase for the next time that similar audio input is encountered.)

Claim 9 is an independent claim that has been amended to include limitations similar to claim 1 and for the reasons set forth above is likewise believed to be patentable and in a condition for allowance.

Claims 2, 3, 5 and 6 depend from claim 1, and for the reasons set forth above are likewise believed to be patentable and in a condition for allowance.

Claim 26 is an independent method claim that has been amended to include limitations similar to those recited above for claim 1. Based on the amendments to claim 26 to include limitations such as: “comparing the audio input to a vocabulary to produce text or a graphic image” and creating “an association between the image, the audio input, the selected location, one of a plurality of different visual notations,” claim 26 is believed to be patentable over the cited art and in a condition for allowance.

Claims 11, 12, 14, 16 and 27-34 have been amended to or did depend from claim 26, and for the reasons set forth above are likewise believed to be patentable and in a condition for allowance.

Claims 35-37 and 39-42 were previously presented. Applicants submit that combination of Rubin-Heck- Mitchell fails to disclose the recited steps of determining since this combination does not distinguish between different annotations with different text and audio associated with each annotation. Since only one annotation or annotations of different colors at best are disclosed by the combination, neither of which can have both audio and/or text associated therewith, Applicants submit that these displaying and retrieving method are unique to the annotation created and provided by the claimed invention and thus are not taught or suggested by the prior art.

## **CONCLUSION**

In sum, Applicants respectfully submit that claims 1-3, 5, 6, 9, 11, 12, 14, 16, 26-37, and 39-42, as presented herein, are patentably distinguishable over all of the art of record. Therefore, Applicants request reconsideration of the basis for the rejections to these claims and request allowance of the claims.

In addition, Applicants respectfully invite Examiner to contact Applicants' representative at the number provided below if Examiner believes it will help expedite furtherance of this application.

Respectfully Submitted,  
GREGORY J. WOLFF, et al.

Date: March 7, 2007

By: /Greg T. Sueoka  
Greg T. Sueoka, Reg. No.: 33,800  
Fenwick & West LLP  
Silicon Valley Center  
801 California Street  
Mountain View, CA 94041  
Tel.: (650) 335-7194  
Fax: (650) 938-5200

20412/06364/DOCS/1701382.1